第65卷 第2期 天 文 学 报 Vol. 65 No. 2
2024年3月 ACTA ASTRONOMICA SINICA Mar., 2024
doi: 10.15940/j.cnki.0001-5245.2024.02.010

快速射电暴观测数据干扰缓解方法研究*

刘艳玲[1,2,3†] 陈卯蒸[1,2,3] 袁建平[1,2] 李 健[1,3] 闫 浩[1,2,3] 翟楠楠[1,2]

(1 中国科学院新疆天文台 乌鲁木齐 830011)
(2 中国科学院大学 北京 100049)
(3 新疆微波技术重点实验室 乌鲁木齐 830011)

摘要 从海量的天文观测数据中快速搜寻罕见的快速射电暴(Fast Radio Burst, FRB)事件, 干扰缓解是其中一项关键而具有挑战的工作. 射频干扰(Radio Frequency Interference, RFI)会淹没真实的天文事件, 还会导致搜寻管线输出大量的假阳性候选体. 由于干扰来源及其种类的复杂性, 目前并没有一种通用的方法可以解决这个问题. 为了降低干扰对FRB观测搜寻的影响, 分析和研究了南山26 m射电望远镜L波段观测数据中的干扰情况, 针对主要的窄带干扰和宽带干扰建立了3层次的干扰缓解处理流程, 从而有效缓解了观测数据的干扰污染情况. 将该流程嵌入到FRB色散动态谱搜寻(Dispersed Dynamic Spectra Search, DDSS)管线中, 实验结果表明, 搜寻管线的检测率和检测精度得到了进一步的提高. 该方法为FRB观测数据干扰缓解处理提供了有价值的参考.

关键词 射电连续谱: 暂现源, 射频干扰, 方法: 数据分析
中图分类号: P161; **文献标识码**: A

1 引言

射电天文学利用射电天文望远镜被动地接收宇宙中极其微弱的射电信号, 而接收到的电磁波信号是非常微弱的, 大概是150–220 dB · W · m^{-2}, 例如, 最强的射电源在300 MHz的功率谱密度大概是10^5 Jy (1 Jy = 10^{-26} W · m^{-2} · Hz^{-1})[1]. 因此, 射电天文观测活动非常容易受到人为活动的无线电干扰, 如移动电话、卫星、飞行器、广播电视以及台址内电子设备等[2]. 另一方面, 随着射电天文技术的发展, 射电望远镜接收机系统灵敏度不断提高, 带宽逐渐增加, 甚至超过射电天文频谱分配保护范围, 而与此同时, 数字接收机、数字终端、商用设备、电气设备、各种线缆的引入以及各种航空航天卫星通信业务的发展使射电天文台址的电磁环境变得更为复杂. 在现在以及未来的射电天文学中, 射频干扰(Radio Frequency Interference, RFI)缓解问题将是主要的挑战之一. 单天线观测最易受到干扰的影响, 因为它们对干扰几乎没有识别力, 会导致天文信号和干扰信号一同进入系统[3].

快速射电暴(Fast Radio Burst, FRB)是一种明亮的、持续时间仅为数毫秒的射电天文瞬变现象[4–6]. 相较于脉冲星短而稳定的脉冲周期特性, 捕捉罕见的FRB就不那么容易. 在无线电瞬变现象和脉冲星搜寻中, 理想化的数据由弱天体物理信号与不相关的高斯噪声叠加组成. 而干扰的出现让情况变得复杂. 干扰会淹没FRB信号, 降低信噪比, 导致检测算法漏检天文事件的发生; 另一方面, 干扰会触发搜寻算法生成假的候选体样本, 增加后期筛选困难. 因此, 对于FRB的搜寻观测, 干扰缓解是一件

2023-02-21收到原稿, 2023-05-04收到修改稿
*国家自然科学基金项目(11903071), 中国科学院“西部之光”人才培养计划(2022-XBQNXZ-015)资助
†liuyanling@xao.ac.cn

非常必要的工作.

射电天文的观测研究一直伴随着RFI缓解的难题. 尽管天文台站一般都建立在远离城市的地方, 但是, 随着城市化建设进程的推进以及人类活动内容和范围的扩大, 台站的电磁环境只会越来越复杂. 图1展示了电磁干扰来源及缓解策略框架示意图. 随着空间无线电技术的发展及台站内电子设备的不断增多, 仅依靠设立无线电静区、管控台站内电子设备等管理方面的努力无法免除射频干扰对天文观测数据的影响. 因此, 必须在接收机链路和数据后处理阶段采取干扰缓解技术措施. 在接收机链路中, 通过在射频模拟前端设计适当的滤波器滤除可能使放大器饱和的强干扰信号[7–8]. 在数字域链路中通过采用实时在线处理和离线后处理方法对更弱的干扰信号进行消减, 具体措施包括时域切除[9–10]、频域切除[11]和空间域切除[12–13]、自适应滤波[14–16]以及机器学习[17]等形式. 另外, 在对数据进行精细分析时, 目测手动标记RFI的离线后处理方式, 依旧是目前常用的方法. 但是随着数据量的增加, 手动干预的方法显得非常笨拙, 特别是对于天体物理瞬变现象的搜寻观测数据, 显得愈加的无能为力. 总体来说, 没有一种通用的方法能够消除所有的干扰, 在射电天文领域干扰缓解是一项复杂的工程化任务.

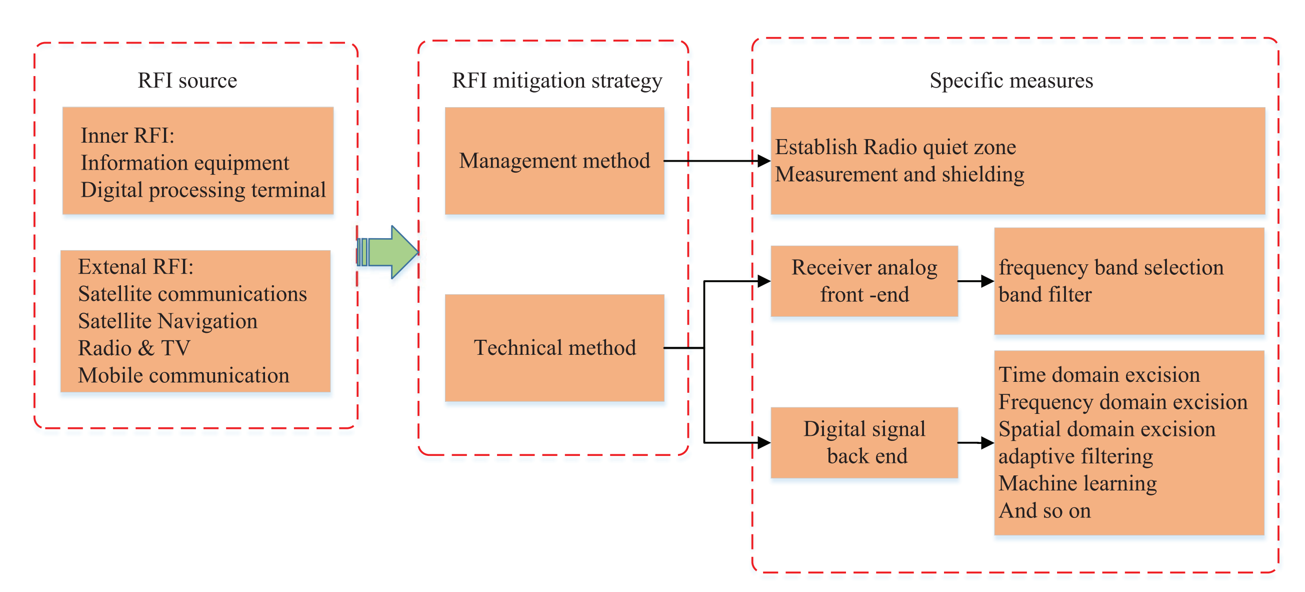


图 1　电磁干扰缓解策略框架示意图

Fig. 1　Schematic diagram of RFI mitigation strategy framework

随着近几年诸如快速射电暴之类的瞬态信号观测研究工作的大量开展, 针对单脉冲观测数据的干扰缓解方法研究也随之而增多. 但是, 由于对实时处理的需求, 使得很多基于峰度、表面拟合或者傅里叶变换等技术的干扰缓解算法因为计算的复杂性而失效. 因此, 对于瞬态信号的处理, 考虑更多的是简单、快速. Men等[18]提出Zero-DM matched filter (ZDMF)处理方法进行瞬时宽带干扰的消减处理, 并同固定窄带干扰消除步骤一同应用到他们自主开发的快速射电暴BEAR (Burst Emission Automatic Roger)搜寻管道中. Boyle等[19]提出了一种边缘阈值分割的干扰抑制算法, 并与时域阈值切除结合, 降低了管道的FRB假阳性检测率. Sclocco等[20]开发了开源的RFIm (RFI Mitigation)库用于执行时域阈值切除和频域阈值切除, 达到消除瞬态观测数据中的宽带干扰和窄带干扰的目的. Morello等[21]开发了四分位距缓解(Inter-Quartile Range Mitigation, IQRM)算法对天体物理瞬态现象和脉冲星搜寻观测数据中的窄带干扰进行处理. 总体来说, 在对快速射电暴观测数据进行干扰缓解

时, 会综合分析观测数据中的干扰情况, 采取多种方法进行处理, 以达到提高数据质量的目的.

本文针对中国科学院新疆天文台南山26 m射电望远镜FRB观测终端产生的观测数据, 分析干扰特点, 选择干扰算法. 通过设计多层次的干扰缓解处理策略对观测数据进行预处理, 为FRB搜寻管线提供高质量的观测数据, 从而提高搜寻管线的检出率与搜寻精度. 在后续章节中, 将详细介绍干扰分类及特点, 多层次的干扰缓解措施实验及应用效果等.

2 信号特征

2.1 FRB信号特征

与脉冲星的单脉冲所观察到的情况类似, 当来自遥远太空的快速射电暴, 穿越星际介质(Intergalactic Medium, IGM)被地基望远镜接收时, 射电脉冲会表现出色散、散射、闪烁等现象, 导致脉冲轮廓的展宽和变形以及强度的变化和频率的漂移. 其中, 色散的影响最为显著, 使得其在不同频率的能量信号到达望远镜的时间有差别, 具体表现为高频信号比低频信号先到达, 即色散效应. 由色散效应引起的到达时间延迟表示为:

$$\Delta t = 4.15 \times 10^6 \times \left(v_2^{-2} - v_1^{-2}\right) \times \mathrm{DM}\,, \qquad (1)$$

其中, 频率v_1, v_2分别为两个频率通道的中心频率, 单位为MHz, DM (Dispersion Measure)为色散量, 定义为电子数密度在视向路径上的积分, 单位为$\mathrm{cm}^{-3}\cdot\mathrm{pc}$, 其表达式如下:

$$\mathrm{DM} = \int_0^d n_{\mathrm{e}} \mathrm{d}l\,, \qquad (2)$$

其中, n_{e}是电子数密度, d为快速射电暴的距离, l为视向路径长度.

色散量是天体物理学中一个重要的测量值, 通过测量天体信号到达时间随频率的变化, 可以得到色散量. 色散量越大, 代表爆发源越远, 通过色散量可以估算源的距离. 在脉冲星或者快速射电暴的观测技术中, 消色散是最重要、最具有特性的技术与方法. 同时, 也是区分来自地面的人为干扰和来自遥远天体的射电信号的主要特征之一.

2.2 RFI信号特征及对FRB的影响

干扰按照其特点有多种分类描述方法, 可以是瞬时脉冲式的或持续性的, 窄带的或宽带的, 周期性的或非周期性的, 台址外部的或内部的等. 而在我们实际观测的数据中, 窄带持续性干扰与瞬时脉冲式干扰叠加的现象是更常见的. FRB一般为瞬时的一次性爆发, 不可预知、鲜少重复. 在目前观测到的700余例快速射电暴中, 重复射电暴只有60多例, 大部分的重复射电暴也只探测到2次. 虽然FRB爆发辐射的能量很高, 但与人为产生的干扰, 如天空的飞机雷达信号、移动通讯、卫星信号及台址电子设备电磁辐射信号相比, 依旧非常微弱, 这使得FRB信号的探测非常困难. 目前针对FRB观测数据, 干扰缓解工作的难点主要在于: 1. 干扰来源复杂, 在观测数据上的表现形态也是多种多样; 2. 干扰是动态变化的, 不同时间、不同频率范围, 干扰表现出很大的差异; 3. 长期看, 干扰情况会持续恶劣. 在射电天文观测研究的路上, 一直伴随着干扰缓解方法的研究, 从未停止过. 针对FRB观测, 虽然我们希望尽量减少数据中的RFI污染情况, 但是, 过度切除RFI会导致数据损失, 甚至可能误将天文事件切除. 另外, 在实时观测搜寻中, 过于复杂的干扰缓解算法或流程势必会增加搜寻管线的处理速度. 因此, 平衡好精度、速度也是FRB搜寻观测中采取干扰缓解措施所需要考虑的问题.

2.3 Peryton的特征

Peryton即佩利冬, 传说中一种半鹿半鸟的怪兽, 而它的影子却为人形. 在射电天文领域, 用它来称呼模仿快速射电暴的地面无线电干扰信号. Parkes 64 m射电望远镜曾观测到这类信号, 后来被鉴定为来自微波炉的干扰[22]. 这类干扰具有色散特征, 很容易被现有的FRB检测算法误判. 但是因为发生的概率很小, 目前现有的搜寻算法并没有对其与FRB事件做进一步的区分. 在本文的研究中, 也并不对此做进一步的研究和区分.

3 南山L波段信号干扰主要来源与分布

新疆天文台南山基地始建于1991年, 目前已拥有26 m射电望远镜系统、40 cm精密光电望远镜系统、1 m直径光学望远镜、多个GPS数据接收系统和太阳色球望远镜等仪器设备, 有多台重要天文观测设备正在建设或者计划建设中. 另外, 人类活动的射电源辐射迅速扩大, 可以预见, 未来台址的电磁环境将会更加复杂. 如图1所示, 射电天文台站电磁干扰来源分为台站内部干扰和台站外部干扰. 台站内部干扰主要来自数字时钟、电源开关、继电器、数字处理设备及生活区电气设备(如无线鼠标、电视机顶盒、笔记本电脑、微波炉等), 其特点是发射信号频率、时间不固定, 位置相对固定. 台站外部干扰主要有移动通信、电视广播、航空导航、航空移动通信、雷达、卫星、业余无线电等, 其特点是发射频段相对固定, 发射功率较大. 天文观测数据中往往掺杂了多种干扰源信息, 仅凭视觉上发现的信号形式突变, 往往很难识别判断干扰来源. 南山L波段接收机主要用于脉冲星和单脉冲搜寻观测, 因此, 本文主要针对L波段的观测数据进行FRB干扰缓解策略的分析与研究. 南山L波段接收机频率范围为1400–1720 MHz, 在这个频段范围内的已知外部干扰见表1.

表 1 1400–1720 MHz频段范围内的外部RFI来源[23]
Table 1 The sources of external RFI in the frequency range of 1400–1720 MHz[23]

RFI Source	Frequency/MHz	Characteristic
GPS	1227.6, 1575.42	Persistent
Inmarsat	1526–1559	Persistent
GLONASS	1602	Persistent
Iridium	1618.85–1626.5	Narrow and broad bands persistent

另外, 从人类社会活动时间特点来看, 在不同的观测时间段, 干扰种类、分布、数量以及强弱也是变化的. 新疆乌鲁木齐本地工作时间是从北京时间上午10点到晚上8点, 干扰状况处于高度活跃的状态, 从晚上8点到第2天上午10点之前, 干扰相对要少很多. 因此, 夜间观测数据中的干扰相对来说会更少一些.

4 干扰缓解方法及处理效果

4.1 观测数据分析

针对FRB搜寻观测数据进行干扰缓解, 一方面是为了提高信噪比, 以便让搜寻算法能够更容易将FRB检测出来, 另一方面, 希望通过减少干扰, 降低搜寻算法误报的概率. 如图2所示, 为新疆天文台南山26 m射电望远镜L波段观测数据的频率-时间强度图. 由于数字终端带宽大于L波段接收机的工作带宽, 因此, 保存的观测数据会有一部分有效观测带宽外信号. 在实际的处理过程中, 如有需要可以将多余的带宽数据删除. 图中主要包含窄带持续性干扰、多个宽带瞬时干扰以及短时窄带干扰. 短时窄带干扰在图2中不容易看出来, 但是在本文后面的标准化处理后的降采样数据中会很明显. 本文将主要针对这3类干扰分层次采取干扰缓解措施.

4.2 干扰缓解实验

4.2.1 窄带持续性干扰消除

我们选用的数据是南山26 m射电望远镜L波段接收机观测的PSR J1935 + 1616频率-时间二维频谱数据, 观测时间为2020年5月3日. L波段接收机工作带宽为320 MHz, 数字终端频率带宽为512 MHz, 时间分辨率为64 μs, 频率分辨率为1 MHz. 在数据处理过程中, 仅保留了接收机工作带宽内的数据. 将二维频谱数据沿时间轴进行样本分割, 每2048次采样为一个样本, 得到的样本为320 × 2048像素的频谱图片. L波段接收机工作带宽范围为1400–1720 MHz, 频带范围内出现的窄带持续干扰频段范围有1450–1465 MHz和1535–1545 MHz. 对受窄带持续干扰污染的频率通道, 通常的处理方法是将受污染的频率通道的数据直接删除. 对于删除的频率通道数据有多种替代方案, 如: (i)用固定常数进行替代; (ii)用剩余通道或部分通道中值的均值替代; (iii)替换为人工高斯噪声.

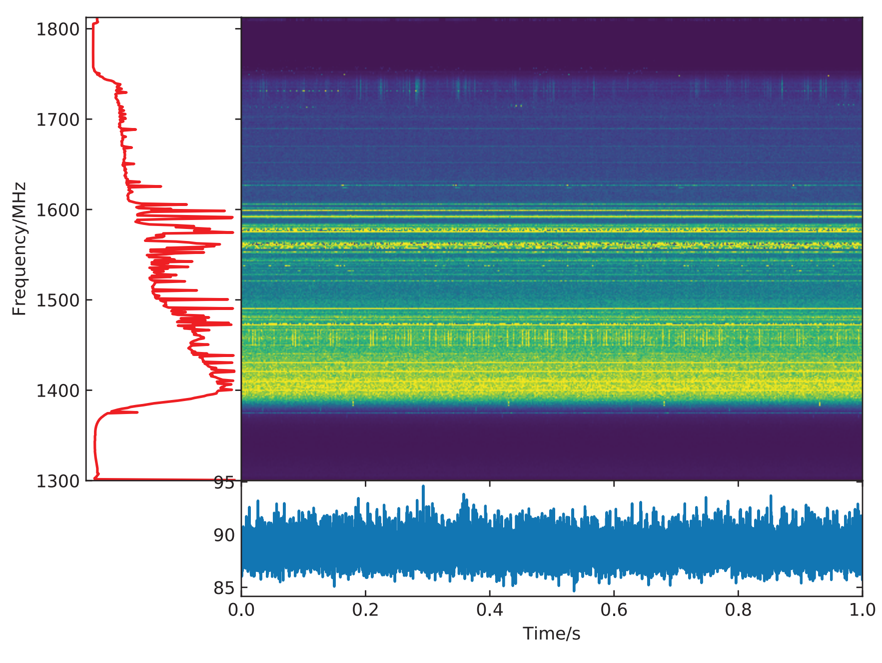


图 2 L波段观测数据动态谱图

Fig. 2 Dynamic spectrum of L-band observation data

消干扰在这里只是作为FRB搜寻管线的数据预处理过程之一, 考虑到后续的过程中, 还需要对数据进行诸如标准化或者归一化之类的处理, 本文选择应用人工高斯噪声进行替换. 另外, 为了提高信噪比, 使数据具有更好的显示度, 我们对数据进行了16倍的降采样处理. 最终得到的待处理样本是320×128像素的频谱图.

4.2.2 窄带短持续干扰消除

相较于4.2.1节的持续性窄带干扰, 此干扰指的是只在部分数据中出现的短持续性窄带干扰. 对于此类干扰, 应用Morello等[21]提出的IQRM算法进行处理. IQRM的目标是在给定的动态谱数据块中, 识别受窄带干扰影响的所有频率通道. 首先是对每个频率通道数据进行频谱统计, 我们采用的是频率通道标准差, 然后再通过计算比较确定哪些通道可能受到了干扰污染. 该算法是非参数的, 并且对序列数据中的趋势具有鲁棒性.

频率通道标准差序列表示为$(x_i)_{i=0}^{n-1}$, 其中x_i代表第i个频率通道的标准差, n为频率通道数. IQRM算法有两个可以设置的参数: 半径r和阈值t. 半径r表示在评估给定数据点的异常状态时, 设定的比较范围. 阈值t表示显著性水平, 代表高斯σ (比较序列的标准差)的倍数, $t\sigma$为判断评估点为异常值的阈值. 具体步骤包括:

首先, 通过输入动态谱数据的频率通道标准差序列, 计算每个频率通道与设定半径范围内的邻近通道标准差的差值

$$\Delta_i^k = x_i - x_{i-k} \, . \tag{3}$$

这里k的取值范围为$[-r, +r]$且$k \neq 0$. 对于得到的序列Δ_i^k用Q_k表示, 采用类似Tukey[24]的异常值规则来识别其中的异常值. 异常值的判断条件如下:

$$\left(\Delta_i^k - \mu_k\right) > t \times \sigma_k \, , \tag{4}$$

其中, μ_k为序列的中位数, 即$\mu_k = Q_k(0.5)$, σ_k为序列的标准差, 用标准化四分位距计算如下:

$$\sigma_k = \frac{\mathrm{IQR}}{1.349} = \frac{Q_3 - Q_1}{1.349} \, , \tag{5}$$

其中, IQR为四分位距, Q_3表示上四分位数, Q_1表示下四分位数.

如果(4)式的条件成立, 则表示x_i的异常程度大于x_{i-k}, 其为异常值. 但这里并不表示x_i应标记为高异常值, 而x_{i-k}为可以忽略的低异常值, 这只是一个相对的比较. 因此, 需要综合多次不同k值计算出的Δ_i, 剔除其中的低异常值, 最后仅留下高异常值, 并对高异常值所对应的频率通道进行干扰缓解. 判断x_i为高异常值的条件是: x_{i-k}在所有轮的计算比较中, (4)式条件成立的次数要少于x_i在所有轮中被判为高异常值的次数.

对于标记出来的频率通道, 本文采用与4.2.1节一致的高斯噪声数据替代方案. 图3为两组处理效果图, 上图为原始观测数据成图. 从图中可以看出, 对窄带干扰处理之后, 单脉冲的图像质量得到了很好的提升.

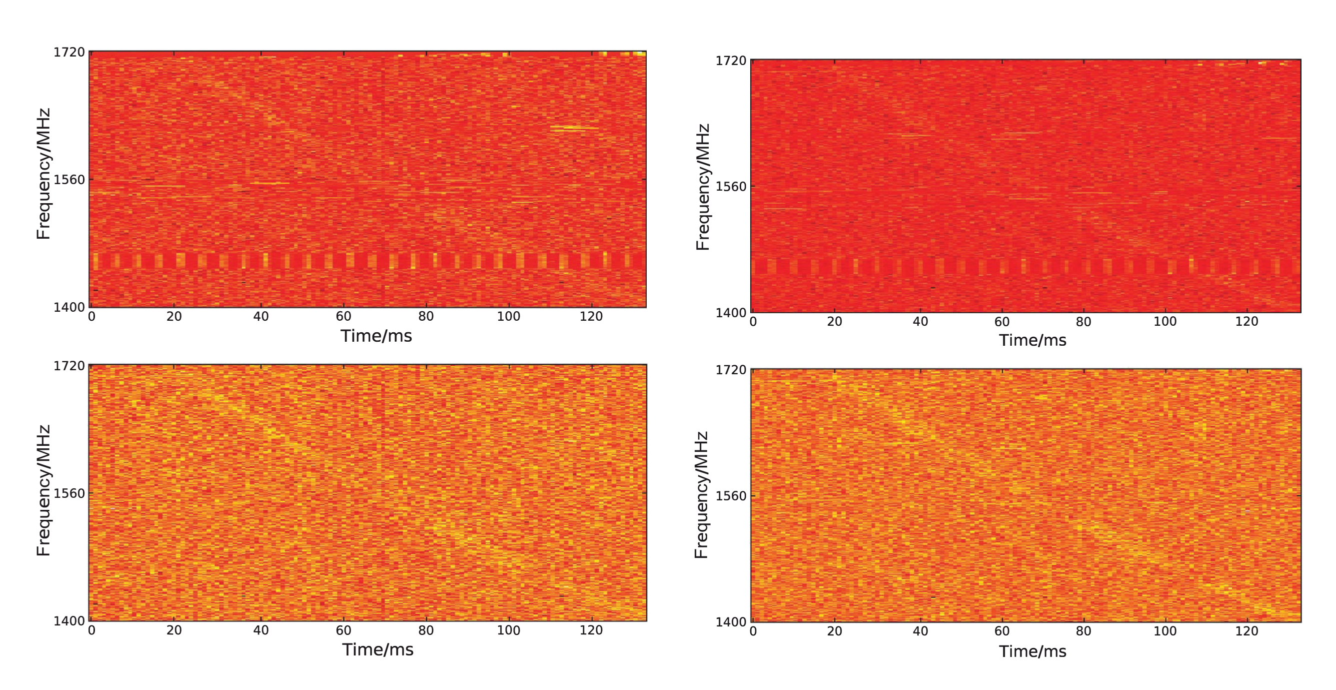


图 3　两例窄带干扰缓解处理结果. 上图为原始频率-时间图, 下图展示了消除窄带干扰后的效果.

Fig. 3　Two examples of narrowband RFI suppression results. The top panels are the raw frequency-time images. The bottom panels show the results of mitigating narrowband RFI.

如果在数据块内的某个频率通道上识别出有短持续性的窄带干扰存在, 通过IQRM算法, 它就会将所在频率通道的数据进行干扰缓解处理. 通常情况下, 窄带持续性干扰也会被准确识别到, 并进行处理. 但是我们做了一个实验, 在原图中的1450–1465 MHz范围的窄带干扰, 如果我们先进行高斯噪声替换, 然后再进行IQRM算法处理, 对于部分图片, 可以进一步提高单脉冲在图片上的显示度. 图4的上图是未经干扰缓解处理的原图, 中图是只应用了IQRM算法去除窄带干扰后的效果, 下图是先对1450–1465 MHz固定窄带干扰进行处理后, 再进行IQRM算法处理后的效果图. 很显然, 加入对固定窄带干扰消除的步骤后, 再应用IQRM算法, 信号质量得到了更显著的提升.

4.2.3　宽带瞬时干扰缓解

宽带瞬时干扰是最为常见且最易引起FRB搜寻软件误判的干扰, 它与FRB事件信号的区别就在于它不具有色散效应. 早在2009年, Eatough等[25]提出了Zero-DM Filter (ZDF)方法进行此类干扰的消除. 该方法通过对每个数据减去该数据所在采样点的所有频率通道数据的均值从而达到消除干扰的目的. 此方法应在消色散处理之前应用, 计算公式如下:

$$S'(i,j) = S(i,j) - \frac{1}{n_{\text{chans}}} \sum_{i=1}^{n_{\text{chans}}} S(i,j)\,, \quad (6)$$

其中, $S(i,j)$为在第i个频率通道, 第j个采样点的数据, n_{chans}表示频率通道数. $S'(i,j)$为消除干扰后, 第i个频率通道, 第j个采样点的数据.

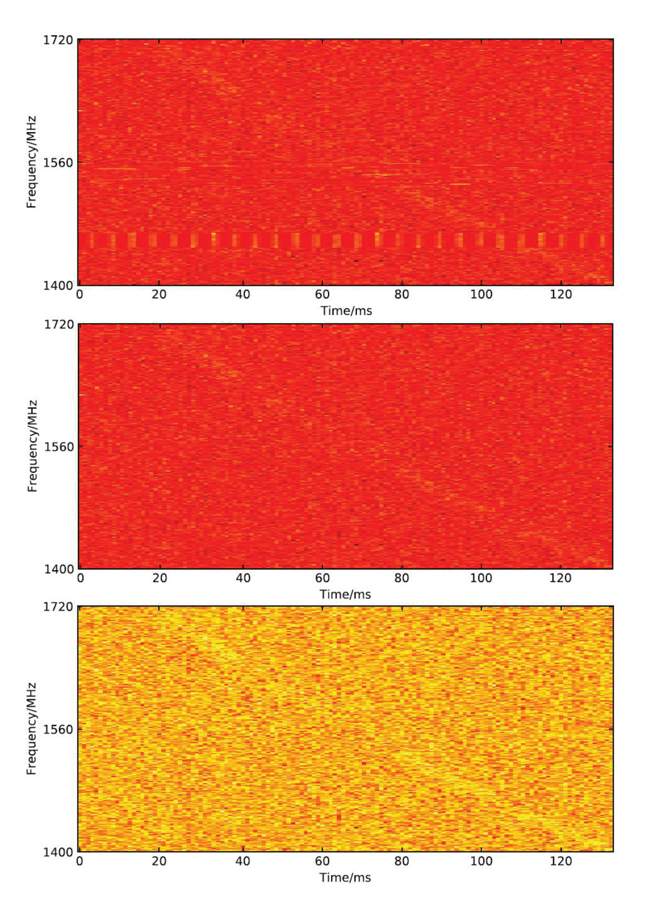


图 4 加入固定窄带干扰缓解处理的效果. 上图是原始频率-时间图, 中图是仅用IQRM算法去除窄带干扰后的效果, 下图是加入固定窄带干扰消除步骤的效果.

Fig. 4 The process effect of adding persistent narrowband RFI mitigation. The top panel is the raw frequency-time image. The middle panel displays the result of mitigating the narrowband RFI with IQRM algrithm only. The bottom panel shows the result of adding the step of removing persistent narrowband RFI.

理论上, 这种方法可以完美消除瞬时宽带干扰. 但是, 在实际的应用中, 会发现宽带干扰在整个频率带宽上的强度并不是均匀一致的, 极端情况是瞬时宽带干扰只出现在部分观测带宽. 因此, 会导致过度切除信号的风险. Men等[18]提出ZDMF处理方法, 该方法是ZDF的升级版本, 它通过计算出每个频率通道对干扰贡献的比例因子, 从而只去除其中对干扰有贡献的部分. 图5为分别用两种方法对同一组数据的处理效果, 其中上图是未经干扰处理的原图, 中图和下图分别是运用ZDF和ZDMF算法进行干扰缓解处理的结果, 可以看出ZDMF的效果更好.

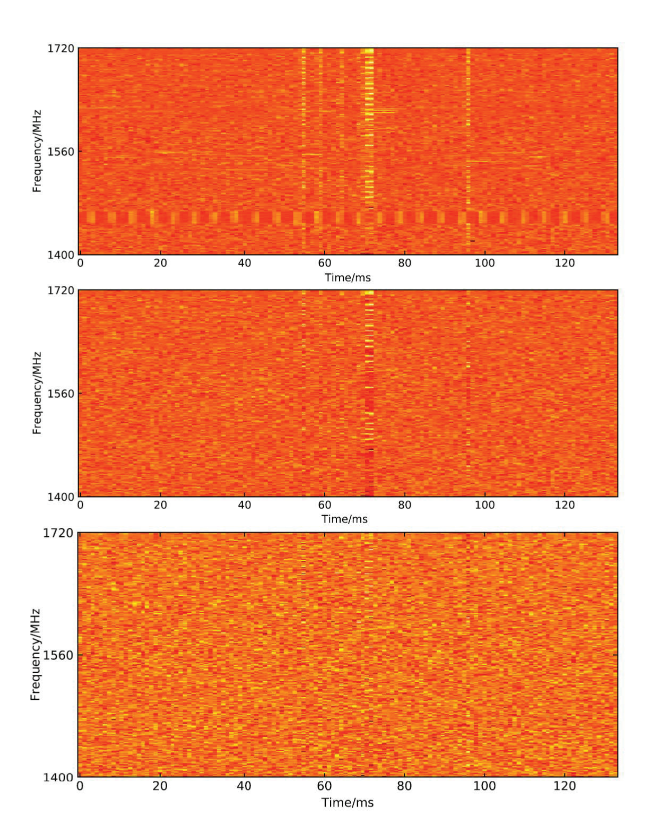


图 5 分别用ZDF和ZDMF技术对同一组数据的干扰缓解处理效果. 上图是原始频率-时间图, 中图是用ZDF进行宽带干扰缓解的效果, 下图是用ZDMF进行宽带干扰缓解的效果

Fig. 5 The results after RFI mitigation using ZDF and ZDMF technique respectively. The top panel is the raw frequency-time image. The middle panel displays the result of mitigating broadband RFI with ZDF technique. The bottom panel shows the result of mitigating broadband RFI with ZDMF technique.

ZDMF方法的总体思想是通过将Zero-DM波形序列拟合到给定频率通道i序列的残差最小, 得到第i个频率通道的基线β_i和比例因子α_i, 残差定义

为:

$$\chi^2 = (S_i - \alpha_i S_{\mathrm{dm}=0} - \beta_i)^2 \ , \tag{7}$$

其中, S_i是第i个频率通道的时间序列. $S_{\mathrm{dm}=0}$是所有频率通道时间序列的平均值, 即Zero-DM波形序列. 比例因子α_i计算如下:

$$\alpha_i = \frac{S_{\mathrm{dm}=0} \cdot S_i - \frac{1}{N}\sum S_i \sum S_{\mathrm{dm}=0}}{S_{\mathrm{dm}=0} \cdot S_{\mathrm{dm}=0} - \frac{1}{N}\sum S_{\mathrm{dm}=0} \sum S_{\mathrm{dm}=0}} \ , \tag{8}$$

这里“·”表示为时间序列的内积, N表示时间序列的数据点数量. 从数据中去除干扰贡献部分的计算公式如下:

$$S_i' = S_i - \alpha_i S_{\mathrm{dm}=0} \ , \tag{9}$$

S_i'即为干扰缓解后的第i个频率通道的数据. 在处理过程中, 我们没有去除偏置β_i, 因为它并不影响FRB的动态谱图样本检测. 针对我们的观测数据, 它的应用效果如图6所示.

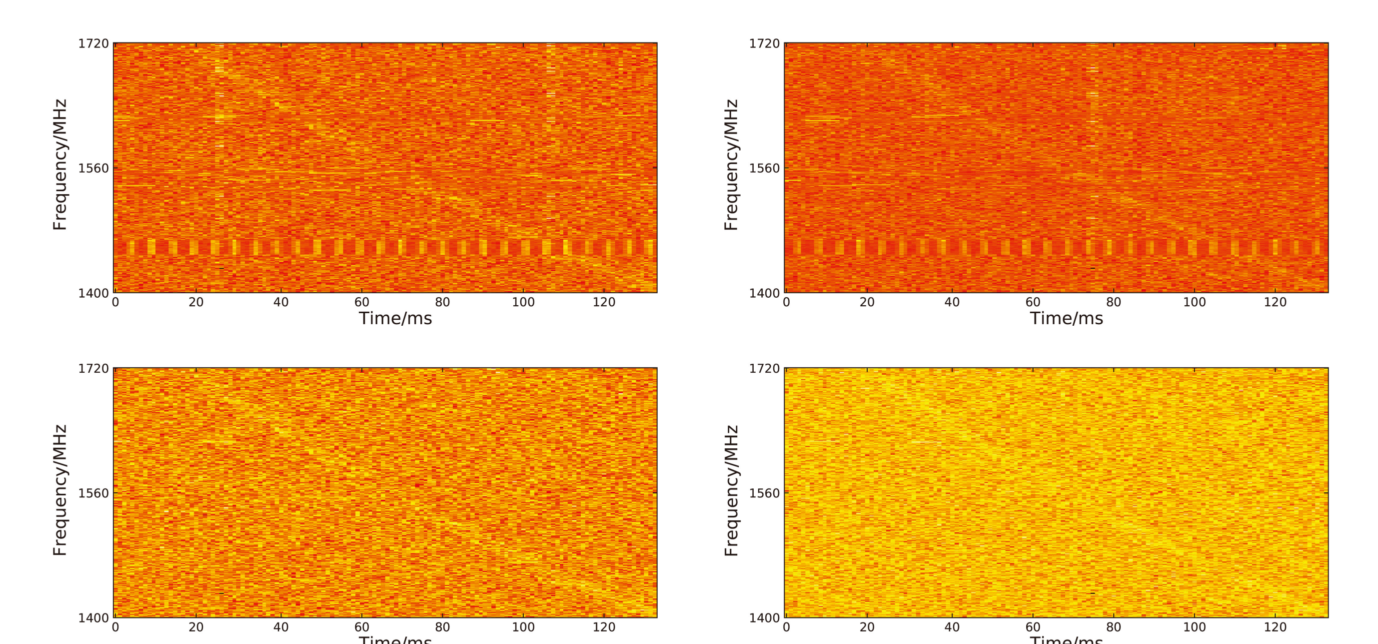


图 6 两例瞬时宽带干扰缓解结果. 上图为原始频率-时间图, 下图是消除瞬时宽带干扰后的效果

Fig. 6 Two examples of instantaneous broadband RFI mitigation results. The top panels are the raw frequency-time images. The bottom panels show the results of mitigating instantaneous broadband RFI.

最后, 为了进一步验证我们设计采取3层干扰缓解策略的必要性, 我们分别单独使用IQRM和ZDMF两种方法与组合使用IQRM+ZDMF干扰缓解流程做了补充对比实验, 实验效果如图7和图8所示, 其中上图是原始数据的频率-时间动态谱图, 中图是仅采用IQRM或者ZDMF处理后的效果图, 下图是采用IQRM+ZDMF组合方式的处理效果图. 可以看出IQRM+ZDMF组合的效果要更优于单独使用IQRM或ZDMF处理的方式.



4.3 处理速度及检测实验

我们将该干扰缓解处理流程应用到开发的FRB DDSS处理管线[26]中. 对一组观测时长为268.44 s的PSR J1935+1616观测数据进行单脉冲搜寻处理, 运行结果如表2所示, 加入ZDMF的干扰缓解算法的检测流程所需要的时间是不加入该缓解方法的10倍以上. 这里需要注意的是, 不同的硬件平台配置、程序语言以及算法设计的执行速度是有很大差别的. 本文提供的测试结果, 只是基于目前的一个单线程的流水处理管线的处理速度, 更

有意义的是以上两种处理流程的对比数据. 另外, 干扰处理流程仅选择对窄带干扰进行处理时, 整个FRB搜寻管线的处理速度是观测数据产生速度的6.8倍以上. 但就目前的搜寻管线而言, 如果想要执行3层次的干扰缓解处理流程, 达到实时搜寻的目的, 那么在干扰缓解处理的环节上就需要进行针对性的设计, 如采取多线程的并行处理方式等以满足速度上的需求. 从单脉冲检测数量上看, 加入ZDMF干扰缓解算法的搜寻管线的单脉冲检出数量要更多一些. 可见, 执行3层次的干扰缓解处理流程后, 数据质量的提高有助于提高搜寻算法的检出率. 这里需要提到的一点是, 虽然我们能够根据PSR J1935+1616的周期计算出数据包中含有的单脉冲数量, 但是这里我们不能用此数据来进一步计算和衡量搜寻管线的检出率, 因为脉冲星的每个单脉冲强度是不一样的, 甚至很多是被淹没在观测背景噪声中的.

表 2　干扰缓解算法执行速度及FRB DDSS (Dispersed Dynamic Spectra Search)搜寻管线的单脉冲检测数

Table 2　The execution speed of RFI mitigation algorithm and pulse detection number produced by the FRB DDSS pipeline

Methods	Processing time/s	Pulse detections
IQRM	39.23	264
IQRM+ZDMF	412.62	287

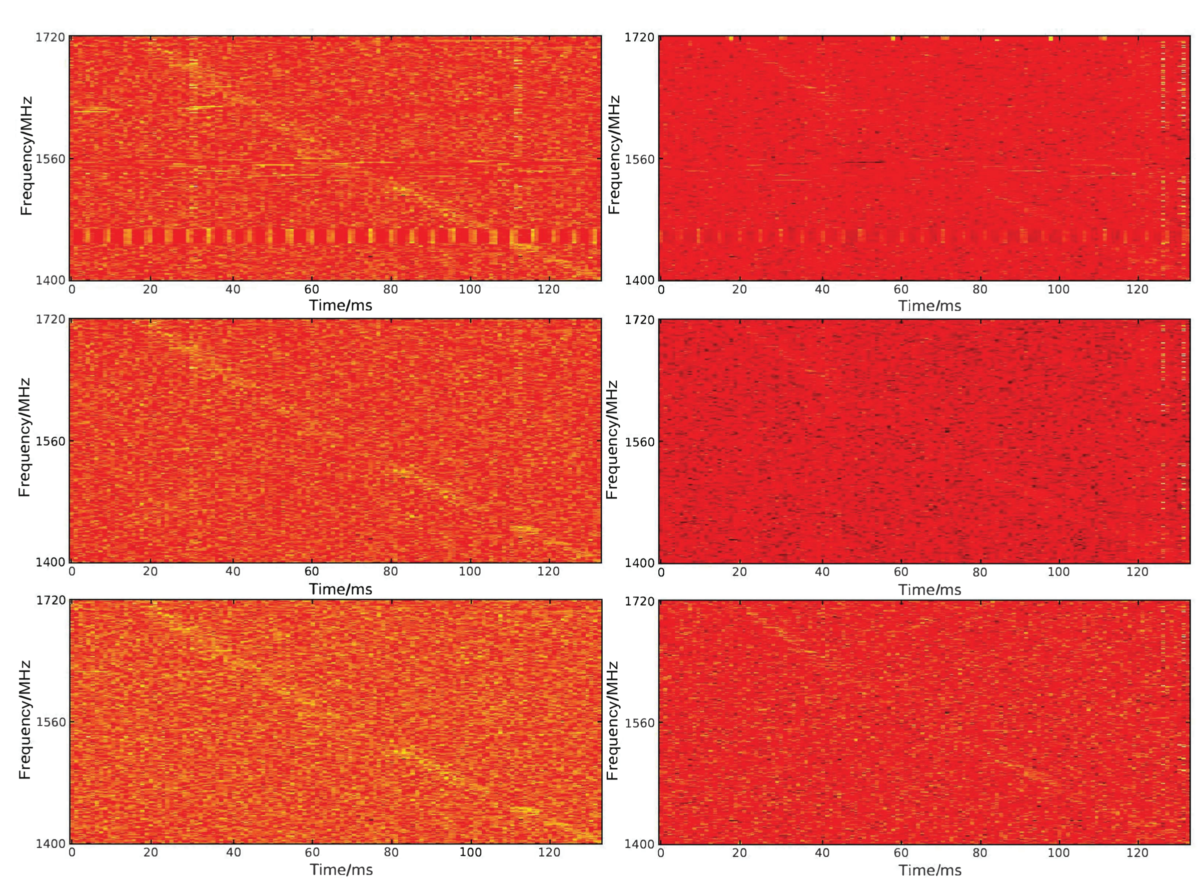


图 7　两例分别用IQRM和IQRM+ZDMF组合的干扰缓解结果. 上图为原始频率-时间图, 中图是用IQRM进行干扰缓解的效果图, 下图是用IQRM+ZDMF组合进行干扰缓解的效果图

Fig. 7　Two examples of the RFI mitigation results using IQRM and IQRM+ZDMF respectively. The top panels are the raw frequency-time images. The middle panels display the results of mitigating RFI with IQRM technique. The bottom panels show the results of mitigating RFI with combining IQRM and ZDMF techniques.

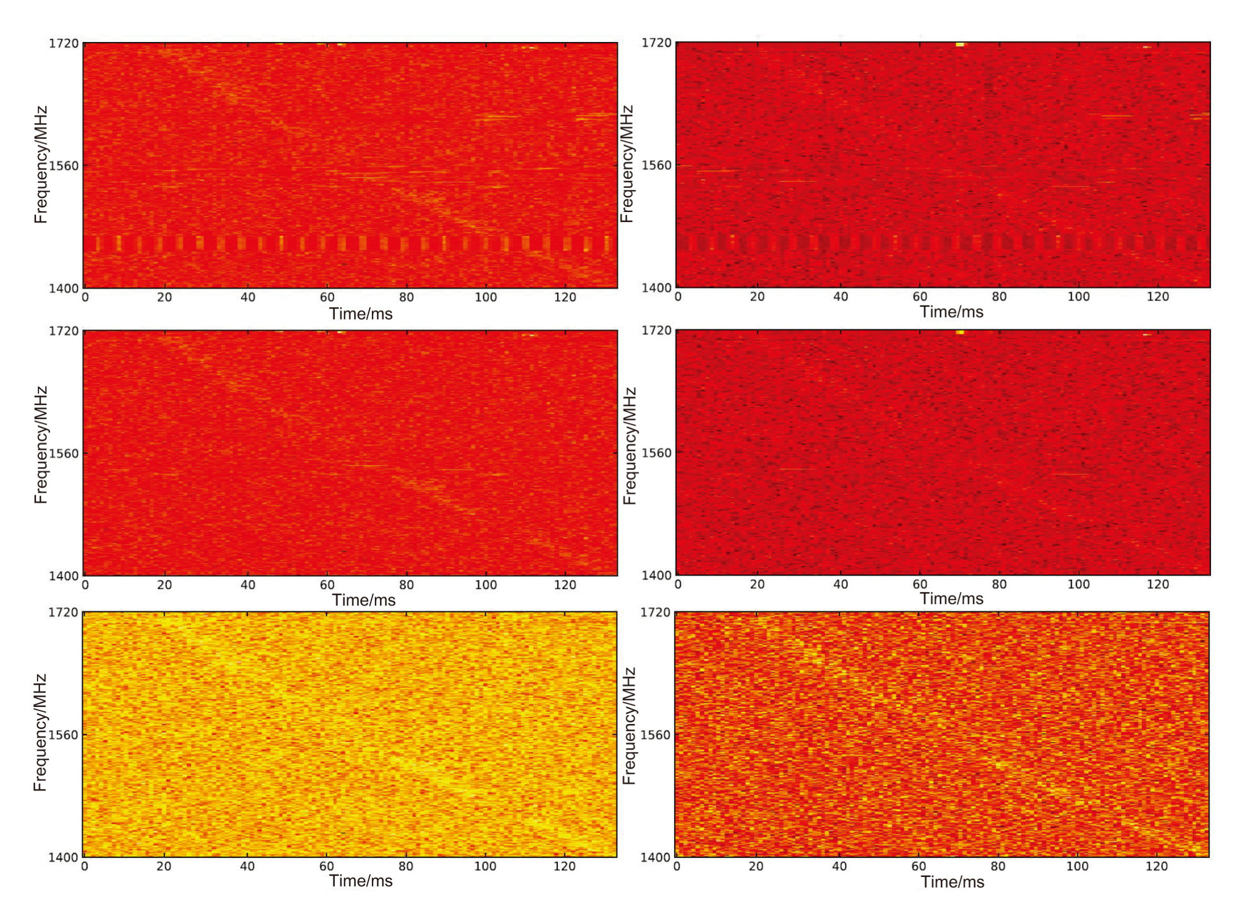


图 8　两例分别用ZDMF和IQRM+ZDMF组合的干扰缓解结果. 上图为原始频率-时间图, 中图是用ZDMF进行干扰缓解的效果图, 下图是用IQRM+ZDMF组合进行干扰缓解的效果图.

Fig. 8　Two examples of the RFI mitigation results using ZDMF and IQRM + ZDMF respectively. The top panels are the raw frequency-time images. The middle panels display the results of mitigating RFI with ZDMF technique. The bottom panels show the results of mitigating RFI with combining IQRM and ZDMF techniques.

5　总结和展望

本文探讨了射电天文观测所面临的干扰环境挑战以及主要的干扰缓解措施, 分析了FRB信号特征及RFI对FRB搜寻观测的影响, 对南山26 m射电望远镜台址电磁干扰环境、干扰主要来源以及分布做了综述. 为了提高FRB检测算法的检出精度, 本文针对南山L波段观测数据中的主要干扰进行了分类, 建立了多层次的干扰缓解处理流程, 实验结果显示, 该流程取得了显著的干扰缓解效果. 但带来的另一个问题是增加了数据处理时间, 在将该干扰缓解流程直接应用到FRB实时搜寻管线中时, 需要提前综合考虑设计实时处理需求. 另外, 该方法具有通用性, 可以为其他射电望远镜的数据处理提供有价值的参考.

致谢 感谢审稿人和编辑老师对文章提出的宝贵建议, 使得文章内容更加的严谨和充实. 感谢新疆天文台南山基地26 m射电望远镜为本文的研究与实验提供了观测数据.

参 考 文 献

[1] Ford J M, Buch K D. Geoscience & Remote Sensing Symposium. Quebec City: IEEE, 2014: 231

[2] Hogden J, Wiel S V, Bower G C, et al. AJ, 2012, 747: 141

[3] Baan W A. General Assembly and Scientific Symposium. Istanbul, Turkey: IEEE, 2011: 1
[4] Lorimer D R, Bailes M, Mclaughlin M A, et al. Science, 2007, 318: 777
[5] Thornton D, Stappers B, Bailes M, et al. Science, 2013, 341: 53
[6] Petroff E, Hessels J W T, Lorimer D R. A&ARv, 2019, 27: 4
[7] Pal S, Lancaster M J, Norrod R D. IEEE Microwave & Wireless Components Letters, 2012, 22: 236
[8] Tuccari G, Caddemi A, Barbarino S, et al. arXiv preprint astro-ph/0501019, 2005
[9] Fridman P A. AJ, 2008, 135: 1810
[10] Lane W M, Cohen A S, Kassim N E, et al. RaSc, 2005, 40: 2189
[11] Nita G M, Gary D E, Liu Z, et al. PASP, 2007, 119: 805.
[12] Nagel J R, Bradley R, Warnick K F, et al. RaSc, 2007, 42: 553
[13] Hellbourg G, Trainini T, Weber R, et al. 2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO). Bucharest, Romania: IEEE, 2012: 200
[14] Ellingson S W, Bunton J D, Bell J F. ApJS, 2001, 135: 87
[15] Poulsen A J, Jeffs B D, Warnick K F, et al. AJ, 2005, 130: 2916
[16] Finger R, Curotto F, Fuentes R, et al. PASP, 2018, 130: 025002
[17] Vos E E, Francois Luus P S, Finlay C J, et al. 2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP). Pittsburgh, PA, USA. IEEE, 2019: 1
[18] Men Y P, Luo R, Chen M Z, et al. MNRAS, 2019, 488: 3957
[19] Boyle J, Sclocco A. 2019 RFI Workshop–Coexisting with Radio Frequency Interference (RFI). Toulouse, France. IEEE, 2019: 24
[20] Sclocco A, Vohl D, van Nieuwpoort R V. 2019 RFI Workshop–Coexisting with Radio Frequency Interference (RFI). Toulouse, France. IEEE, 2019: 1
[21] Morello V, Rajwade K M, Stappers B W. MNRAS, 2021, 510: 1393
[22] Petroff E, Keane E F, Barr E D, et al. MNRAS, 2015, 451: 3933
[23] 安涛, 陈骁, Mohan P, 等. 天文学报, 2017, 58: 20
[24] Tukey J W. Exploratory Data Analysis. Boston, US: Addison-Wesley, 1977: 39-56
[25] Eatough R P, Keane E F, Lyne A G. MNRAS, 2009, 395: 410
[26] Liu Y L, Li J, Liu Z Y, et al. RAA, 2022, 22: 15289

Radio Frequency Interference Mitigation Methods for Fast Radio Burst Observational Data

LIU Yan-ling[1,2,3] CHEN Mao-zheng[1,2,3] YUAN Jian-ping[1,2] LI Jian[1,3] YAN Hao[1,2,3] ZHAI Nan-nan[1,2]

(1 Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi 830011)
(2 University of Chinese Academy of Sciences, Beijing 100049)
(3 Xinjiang Key Laboratory of Microwave Technology, Urumqi 830011)

Abstract To quickly search for rare fast radio bursts (FRBs) from massive astronomical observation data, Radio Frequency Interference (RFI) mitigation is one of the crutial and challenging tasks. RFI will produce a large number of false positive candidates, and even submerge real astronomical events. Due to the complexity of the sources and types of RFI, there is currently no universal method to solve this problem. In order to reduce the impact of RFI on FRB search, the RFI in L band observational data of the Nanshan 26 m radio telescope was analyzed and studied. A three-level RFI mitigation procedure was established for the main narrowband RFI and broadband RFI, which effectively alleviated the RFI pollution of observational data. Embedding this procedure into the FRB DDSS (Dispersion Dynamic Spectra Search) pipeline, experimental results show that the detection rate and accuracy of the search algorithm have been further improved. This method provides valuable reference for RFI mitigation of FRB observational data.

Key words radio continuum: transients, radio frequency interference (RFI), methods: data analysis